Available online at www.sciencedirect.com
SCIENCE DIRECT®

Advances in Applied Mathematics 33 (2004) 136–145

ADVANCES IN Applied Mathematics

www.elsevier.com/locate/yaama

The rapid mixing of random walks defined by an n-cube

Haiyan Chen * and Fuji Zhang

Department of Mathematics, Xiamen University, Xiamen 361005, China

Received 5 March 2003; accepted 18 August 2003

Available online 17 January 2004

Abstract

Inspired by the mutation operator in genetic algorithms, we construct a complete weighted graph G from the n-cube Q_n. The eigenvalues and conductance of G are determined first, then we show the rapid mixing of the random walk on G.

Keywords: Random walks; Markov chain; Conductance; Rapid mixing

1. Introduction

A *random walk* on a graph is a walk $X_0X_1\cdots$ obtained in a certain random fashion. A (discrete-time) *Markov chain* on a finite set V of states is a sequence of random variables $X_0,\ X_1,\ \ldots$ taking values in V such that for all $x_0,\ldots,x_{t+1}\in V$, the probability of $X_{t+1}=x_{t+1}$, conditional on $X_0=x_0,\ldots,X_t=x_t$, depends only on X_t. In fact, there is not much difference between the theory of random walks on graphs and the theory of finite Markov chains, see Kemeny and Snell [10]; every Markov chain can be viewed as a random walk on a directed weighted graph. Similarly, time-reversible Markov chains can be viewed as random walks on undirected weighted graphs. In this paper, the weighted graph is connected and undirected, so the random walk on it is a reversible, irreducible Markov chain.

The classical theory of random walks deals with random walks on simple, but infinite graphs, like grids, and studies their qualitative behaviour: does the random walk return to its starting point with probability one? does it return infinitely often? how to determine

* Corresponding author.
E-mail address: chey6@sohu.com (H. Chen).

0196-8858/$ – see front matter
doi:10.1016/j.aam.2003.08.001

its hitting probability? For more details, see Spitzer [14] or Doyle and Snell [4] or Thomassen [15]. But finite graphs have received much attention recently, and the aspects studied are more quantitative: how long we have to walk before we return to the starting point? before we hit a given point? how fast does the distribution of the walking point tend to its stationary distribution?

Much of the recent interest in random walks is motivated by important algorithmic applications [9]. Random walks can be used to reach “obscure” parts of large sets, and also to generate random elements in large and complicated sets, such as the set of lattice points in a convex body or the set of perfect matching in a graph (which, in turn, can be used for the asymptotic enumeration of these objects). In those applications, to show that the random walk “mixes rapidly”—i.e., is close to the stationary distribution after a number of steps which is bound above by a polynomial input is often the important and challenging part. So far, various methods have been developed to prove the rapid mixing of a random walk: eigenvalues, coupling, conductance, strong stopping times, etc. [2]. In this paper, we use eigenvalues and conductance to prove the rapid mixing of random walks on a special kind of weighted graph. Incidently, we point out, in general, the computation of the eigenvalues is difficult because the underlying graph is exponentially large. The conductance, a combinatorial quantity introduced by Jerrum and Sinclair [8], on one hand, has turned out to be of great use in practice, on the other hand, the conductance itself is not an easy parameter to handle. It is NP-hard to determine it even for an explicitly given random walk. However, for the special random walk in this paper, we can not only compute the eigenvalues but also the conductance explicitly.

The remainder of the paper is organized as follows. In Section 2, we introduce some definitions and notations, and give three theorems we need in the following sections. In Sections 3, we mainly compute the eigenvalues of transition probability matrix P. In Section 4, we derive the conductance and prove the rapid mixing of the random walk.

2. Definitions and notations

Let Q_n denote the n-cube, that is, $V(Q_n) = \{(x_i)_{i=1}^n \mid x_i = 0 \text{ or } 1\}$, two vertices are adjacent if and only if they differ in exactly one coordinate. It is known, Q_n is not only a vertex-transitive graph but a distance-regular graph with 2^n vertices and diameter n [5]. From Q_n, we construct a complete weighted graph G with one loop attaching to each vertex such that $V(G) = V(Q_n)$ and the weight of edge uv is $\omega(u, v) = p^d(1-p)^{n-d}$, where $d = d_{Q_n}(u, v)$ is the distance between u and v in Q_n, p (< 1) is a fixed positive number. Now, for every vertex $u \in V(G)$, let $\omega_u = \sum_{v\in\Gamma(u)} \omega(u, v)$, and for $u, v \in V(G)$, define

$$p_{uv} = \frac{\omega(u, v)}{\omega_u}.$$

Thus $P = (p_{uv})$ is an $n \times n$ doubly-stochastic matrix. In the following, by a random walk (RW) on G we shall mean a RW with P as its transition probability matrix, in other words, p_{uv} is the probability of going from u to v. Let k_r be the number of vertices at distance r

from a fixed vertex u in $V(Q_n)$, then $k_r = \binom{n}{r}$ is independent of u since Q_n is a distance-regular graph, furthermore, we can easily obtain some facts about the RW on G defined above:

(i) for any $u \in V(G)$, $\omega_u = \sum_{r=0}^{n} k_r p^r (1-p)^{n-r} = 1$;
(ii) the RW on G is a reversible, aperiodic and irreducible Markov chain with $\pi = (1/2^n, 1/2^n, \ldots, 1/2^n)$ as its stationary distribution.

Thus let $\mathbf{p}_0$ be the initial probability distribution on $V(G)$, and $\mathbf{p}_t = \mathbf{p}_0 P^t$ the distribution of X_t. Then $\mathbf{p}_t$ tends to the stationary distribution π. Our main aim in this paper is to judge whether the RW on G is rapid mixing. More precisely, a RW on a weighted graph G of order N is *rapidly mixing* if there is a polynomial f such that if $0 < \varepsilon < 1/3$ and $t \geqslant f(\log N)\log(1/\varepsilon)$, then $d_1(t) = \sum_i |\mathbf{p}_t(i) - \pi(i)| \leqslant \varepsilon$. To this end, just as we have pointed out in the introduction, we will use eigenvalues and conductance. So let us first give the relationships between them and rapid mixing. For notational simplicity, we take $V(G) = \{1, 2, \ldots, N\}$. Let $\lambda_1 \geqslant \lambda_2 \geqslant \cdots \geqslant \lambda_N$ be the eigenvalues of the transition probability matrix P, enumerated with multiplicities, we know from linear algebra that $1 = \lambda_1 > \lambda_2 \geqslant \lambda_3 \geqslant \cdots \geqslant \lambda_N \geqslant -1$. Then we have

Theorem 2.1 [9]. *For a finite reversible Markov chain, with $\pi_0 = \min_i \pi(i)$, for any t,*

$$d_1(t) = \sum_i \bigl|p_t(i) - \pi(i)\bigr| \leqslant \frac{1}{\pi_0}\bigl[\max\bigl(\lambda_2, |\lambda_N|\bigr)\bigr]^t.$$

In fact, the crucial quantity here is λ_2, rather than $|\lambda_N|$. This can be made sure by "slowing down" our random walk, as we shall do below. Let $P' = (I + P)/2$ be the new transition probability matrix. Then this new random walk, named the *lazy random walk* (LRW), stays at the current state with probability more than $1/2$ and ensures all eigenvalues of P' nonnegative, while converging at least about half as fast as the old one. So the rapid mixing of RW on G is equivalent to the rapid mixing of LRW on G. For the LRW, it is easy to see that $\lambda_N \geqslant 0$ and the maximum is λ_2 [3,9]. Then from the definition of rapid mixing and Theorem 2.1, if λ_2 is bound away from 1, then the LRW is rapid mixing, so is the RW. As to the other concept we will discuss, *the conductance*, defined as

$$\Phi = \min_{\pi(U) \leqslant 1/2} \frac{\sum_{i \in U} \sum_{j \in V \setminus U} \pi(i) p(i,j)}{\pi(U)},$$

where $\pi(U) = \sum_{i \in U} \pi(i)$, $U \subseteq V$.

Theorem 2.2 [11]. *Every LRW on a weighed graph G of order N with conductance Φ satisfies*

$$d_1(t) = \sum_i \bigl|p_t(i) - \pi(i)\bigr| \leqslant (2N)^{1/2}\bigl(1 - \Phi^2\bigr)^{t/2}.$$

We observed that if the conductance of LRW is bound away from zero, then the LRW is rapid mixing. So the main work in the following is to compute the eigenvalues and conductance, but before that we will give one more theorem we need below.

Theorem 2.3 [1,13]. *For a LRW on a weighted graph G with conductance Φ, we have*

$$1 - 2\Phi \leqslant \lambda_2 \leqslant 1 - \frac{1}{2}\Phi^2.$$

3. The eigenvalues of P

Let A be the adjacency matrix of Q_n, A_i $(i = 0, 1, \ldots, n)$ be the ith distance matrix of G, that is,

$$A_i(u, v) = \begin{cases} 1, & \text{if } d_{Q_n}(u, v) = i; \\ 0, & \text{otherwise.} \end{cases}$$

Then A_i is a polynomial of degree i in A_1 $(= A)$ for all i since Q_n is the distance-regular graph [5]. If the polynomial denoted by $p_i(x)$, then $A_i = p_i(A)$, $i = 0, 1, \ldots, n$. Let G be the weighted graph, and P be the transition probability matrix of the RW on G defined as in Section 2, we have

Lemma 3.1. $P = \sum_{i=0}^{n} p^i (1-p)^{n-i} A_i$.

Proof. In view of the definition of G, the weight of uv only depends on the distance of u, v in Q_n. $\square$

From this lemma we see the eigenvalues of P can be expressed by the eigenvalues of A. To compute the eigenvalues of P explicitly, we must do more preparative work. Fix a vertex u in Q_n, let $s_r = \{v \mid d(u, v) = r\}$, then

$$a_r = \left|\Gamma(v) \cap s_r\right|, \qquad b_r = \left|\Gamma(v) \cap s_{r+1}\right|, \qquad c_r = \left|\Gamma(v) \cap s_{r-1}\right|$$

are independent of the choice of v in s_r. It is clearly that $b_0 = n$, $c_1 = 1$, and $a_0 = c_0 = b_n = 0$, furthermore we have the following lemma:

Lemma 3.2. *For* Q_n, $a_r = 0$, $c_r = r$, $b_r = n - r$ $(r = 0, 1, \ldots, n)$.

Proof. $a_r = 0$ is trivial.

For c_r, b_r, first by counting the edges joining a vertex in $s_{r-1}(u)$ to a vertex in $s_r(u)$, we have the equality

$$k_{r-1} b_{r-1} = k_r c_r, \tag{3.1}$$

where $k_r = |s_r| = \binom{n}{r}$. So

$$\frac{c_r}{b_{r-1}} = \frac{r}{n-(r-1)}. \tag{3.2}$$

By (3.2) and $c_0 = 1,\ c_1 = 1,\ b_n = 0,\ b_0 = n$ we obtain

$$c_r = r, \quad b_r = n - r, \quad \text{for } r = 0, 1, \ldots, n. \qquad \square$$

Lemma 3.3 [5]. *Let G be a distance-regular graph. Then*

$$A_1 A_i = b_{i-1} A_{i-1} + a_i A_i + c_{i+1} A_{i+1}.$$

From Lemmas 3.2 and 3.3, for Q_n, we have

$$x p_i(x) = \bigl(n - (i-1)\bigr) p_{i-1}(x) + (i+1) p_{i+1}(x),$$

and thus

$$i p_i(x) = x p_{i-1}(x) - \bigl(n - (i-2)\bigr) p_{i-2}(x). \tag{3.3}$$

Let $\{p_i(x)\}_{i \geqslant 0}$ be a sequence of polynomials determined by recurrence-relation (3.3), where $p_0(x) = 1,\ p_1(x) = x$. Let $f(x,t)$ denotes the generating function of $\{p_i(x)\}_{i \geqslant 0}$, that is

$$f(x,t) = \sum_{i \geqslant 0} p_i(x) t^i. \tag{3.4}$$

Then we have the following theorem:

Theorem 3.4. $f(x,t) = (1+t)^{\frac{n+x}{2}} (1-t)^{\frac{n-x}{2}}$.

Proof. By (3.4)

$$\frac{\partial f(x,t)}{\partial t} = \sum_{i \geqslant 1} i p_i(x) t^{i-1}.$$

Substituting (3.3) to the right-hand side of this equality, we have

$$\begin{aligned}
\frac{\partial f(x,t)}{\partial t} &= \sum_{i \geqslant 2} \Bigl(x p_{i-1}(x) - \bigl(n - (i-2)\bigr) p_{i-2}(x)\Bigr) t^{i-1} + x \\
&= x \sum_{i \geqslant 1} p_{i-1}(x) t^{i-1} - n \sum_{i \geqslant 2} p_{i-2}(x) t^{i-1} + \sum_{i \geqslant 3} (i-2) p_{i-2}(x) t^{i-1} \\
&= x f(x,t) - n t f(x,t) + t^2 \frac{\partial f(x,t)}{\partial t}.
\end{aligned}$$

Whence

$$\frac{\partial f(x,t)}{\partial t} = \frac{x-nt}{1-t^2} f(x,t). \tag{3.5}$$

The general solution of (3.5) is

$$\begin{aligned} f(x,t) &= c\exp^{\int \frac{x-nt}{1-t^2}\,\mathrm{d}t} = c\exp^{\log(1-t^2)^{n/2} + x\log(\frac{1+t}{1-t})^{1/2}} \\ &= c(1+t)^{\frac{n+x}{2}}(1-t)^{\frac{n-x}{2}}. \end{aligned}$$

Since $p_0(x) = 1$, so the constant $c = 1$, the result is obtained. □

Lemma 3.5 [12]. *The eigenvalues of Q_n are $n-2k$, $k = 0,\dots,n$, the multiplicity is $\binom{n}{k}$.*

After the preparation we have done, we are ready to prove one of the main results in this paper.

Theorem 3.6. *If P is the transition probability matrix of RW on G, then its eigenvalues are $(1-2p)^k$, $k = 0,1,\dots,n$. The multiplicity is $\binom{n}{k}$.*

Proof. By definition

$$\begin{aligned} P &= (1-p)^n A_0 + p(1-p)^{n-1} A_1 + \cdots + p^n A_n \\ &= (1-p)^n I + p(1-p)^{n-1} p_1(A) + \cdots + p^n p_n(A). \end{aligned}$$

Then the eigenvalues of P are

$$(1-p)^n + p(1-p)^{n-1} p_1(\lambda) + \cdots + p^n p_n(\lambda),$$

where λ are the eigenvalues of A. Let $p/(1-p) = t$, then

$$\begin{aligned} &(1-p)^n + p(1-p)^{n-1} p_1(\lambda) + \cdots + p^n p_n(\lambda) \\ &\quad = (1-p)^n \big(1 + p_1(\lambda)t + \cdots + p_n(\lambda)t^n\big) \\ &\quad = (1-p)^n \sum_{i+j=0}^{n} (-1)^j \binom{\frac{n+\lambda}{2}}{i}\binom{\frac{n-\lambda}{2}}{j} t^{i+j}. \end{aligned}$$

Using $n-2k$ to substitute λ, we obtain

$$(1-p)^n \sum_{i=0}^{n} p_i(n-2k)\left(\frac{p}{1-p}\right)^i$$

$$= (1-p)^n \sum_{i+j=0}^{n} (-1)^j \binom{n-k}{i} \binom{k}{j} \left(\frac{p}{1-p}\right)^{i+j}$$

$$= (1-p)^n \left(\frac{1}{1-p}\right)^{n-k} \left(\frac{1-2p}{1-p}\right)^k = (1-2p)^k.$$

Apparently the multiplicity is $\binom{n}{k}$. □

4. Conductance and rapid mixing

In this section, we shall show that the RW on G is rapidly mixing by showing that the LRW on G is rapidly mixing. First from Theorem 3.6, we immediately derive the following theorem:

Theorem 4.1. *Let P' be the transition probability matrix of the LRW on G. Then the eigenvalues of P' are*

$$\frac{1+(1-2p)^k}{2}, \quad k = 0, 1, \ldots, n,$$

and the multiplicity is $\binom{n}{k}$. Furthermore, the second-largest eigenvalue of P' is

$$\lambda_2 = \begin{cases} 1-p, & \text{if } p \leqslant 1/2; \\ 1-2p+2p^2, & \text{if } p > 1/2. \end{cases}$$

Proof. By the definition of LRW on G, $P' = (I+P)/2$, from Theorem 3.6, the results are trivial. □

Now we will turn to compute the conductance of the LRW on G by its definition and relationship with eigenvalues of P' showed by Theorem 2.3. By definition the conductance

$$\Phi' = \min_{|U| \leqslant 2^{n-1}} \sum_{i\in U, j\in \overline{U}} p'(i,j)/|U| = \frac{1}{2} \min_{|U| \leqslant 2^{n-1}} \frac{\sum_{i\in U, j\in \overline{U}} p(i,j)}{|U|} = \frac{1}{2}\Phi.$$

Theorem 4.2. *For LRW on G, we have*

$$\Phi' = \begin{cases} p/2, & \text{if } 0 < p \leqslant 1/2; \\ p(1-p), & \text{if } 1/2 < p < 1. \end{cases}$$

Proof. (i) For $0 < p \leqslant 1/2$, first by Theorems 2.3 and 4.1, we obtain

$$1 - p \geqslant 1 - 2\Phi',$$

whence

$$\Phi' \geqslant \frac{p}{2}. \tag{4.1}$$

On the other hand, let us consider the subset $U = \{(x_i) \in V(G)\colon x_1 = 1\}$ and $\overline{U} = \{(x_i) \in V(G)\colon x_1 = 0\}$, clearly $|U| = 2^{n-1}$. Since Q_n is vertex-transitive, we have

$$\begin{aligned}\sum_{i\in U, j\in\overline{U}} \frac{p(i,j)}{|U|} &= p(1-p)^{n-1} + \binom{n-1}{1} p^2(1-p)^{n-2} + \cdots + \binom{n-1}{n-1} p^n \\ &= p\left((1-p)^{n-1} + \binom{n-1}{1} p(1-p)^{n-2} + \cdots + \binom{n-1}{n-1} p^{n-1}\right) \\ &= p.\end{aligned}$$

In fact, for any $i \in U$, there are $\binom{n-1}{k}$ vertices in $\overline{U}$ with distance $k+1$ from i, hence by definition

$$\Phi' \leqslant \frac{p}{2}. \tag{4.2}$$

By (4.1) and (4.2) we have

$$\Phi' = \frac{p}{2}, \quad \text{if } 0 < p \leqslant 1/2.$$

(ii) For $1/2 < p < 1$, first also by Theorems 2.3 and 4.1, we obtain

$$1 - 2\Phi' \leqslant 1 - 2p + 2p^2,$$

whence

$$\Phi' \geqslant p(1-p). \tag{4.3}$$

On the other hand, let us consider the subset $U = \{(x_i) \in V(G)\colon x_1 = x_2\}$ and $\overline{U} = \{(x_i) \in V(G)\colon x_1 \neq x_2\}$, clearly $|U| = 2^{n-1}$. Since Q_n is vertex-transitive, we have

$$\begin{aligned}\sum_{i\in U, j\in\overline{U}} \frac{p(i,j)}{|U|} &= 2p(1-p)\left((1-p)^{n-2} + \binom{n-2}{1} p(1-p)^{n-3} + \cdots + \binom{n-2}{n-2} p^{n-2}\right) \\ &= 2p(1-p).\end{aligned}$$

In fact, for every $i \in U$, there are $2\binom{n-2}{k}$ vertices in $\overline{U}$ with distance $k+1$ from i, hence by definition

$$\Phi' \leqslant p(1-p). \tag{4.4}$$

From above, we have

$$\Phi' = p(1-p). \qquad \square$$

From the procedure of proving the above theorem, we see the lower bound of Theorem 2.3 is tight, but in general it is not the case. So the method we use to compute the conductance is not suitable to RW on a general graph.

From Theorems 4.1 and 4.2, we observe that the second-largest eigenvalue and the conductance of LRW on G only depend on parameter p. So by Theorems 2.1 or 2.2, we immediately obtain the aim of this paper.

Theorem 4.3. *The LRW on G is rapid mixing, so the RW on G.*

Remark 1. We deal with the rapid mixing of RW on this kind of weighted graph mainly motivated by mutation operator in standard genetic algorithms. Clearly the RW on G defined in this paper acts just as the mutation operator acts on $V(G)$ (the individual space in standard genetic algorithms), p as the mutation rate. So from above results, we derive that the mutation operator in standard genetic algorithm is rapidly mixing. Furthermore, when the mutation rate $0 < p \leqslant 1/2$, the convergent speed of the mutation operator increases with p; while when the mutation rate $1/2 < p < 1$, the convergent speed decreases as p increases.

Remark 2. Although the work in this paper was motivated by mutation operator in standard genetic algorithms, one can see that the Markov chain considered in this paper is the strong product of the two-point space. (The "strong product" of Markov chains has underlying state space given by the Cartesian product, but with transition probability given by $p(x, y) = \prod_{i=1}^{n} p(x_i, y_i)$, for strictly definition, see [6,7].) As a rough heuristic it seems that spectral values and conductance of a Markov chain on the strong product X_s^n should be around n times larger than those of the regular cross product, since all n terms change simultaneously, and perhaps $\lambda_2(X_s^n) = \lambda_2(X)$ or $\Phi(X_s^n) = \Phi(X)$. But in [7], the authors pointed out, in general the strong product does not have this property because the strong product of non-lazy two points space does not. Then how about the case of lazy Markov chain? Theorems 4.1 and 4.2 in this paper show that the strong product of a lazy two point space (i.e., $0 < p \leqslant 1/2$) satisfies $\lambda_2(X_s^n) = \lambda_2(X)$ and $\Phi(X_s^n) = \Phi(X)$, while the strong product of a non-lazy two points space (i.e., $1/2 < p < 1$) does not.

Acknowledgment

We thank the referee for his helpful comments, especially he provided the content of Remark 2.

References

[1] N. Alon, V.D. Milman, λ_1, isoperimetric inequalities for graphs and super concentrators, J. Combin. Theory Ser. B (1985) 73–88.
[2] E. Behrends, Introduction to Markov Chain: with Special Emphasis on Rapid Mixing, Vieweg-Verlag, 1999, 220 seiten.
[3] B. Bollobás, Modern Graph Theory, Springer-Verlag, New York, 1998.
[4] P.G. Doyle, J.L. Snell, Random Walks and Electrical Networks, Mathematical Association of America, Washington, DC, 1984.
[5] C.D. Godsil, Algebraic Combinatorics, Chapman and Hall, New York, 1993.
[6] C. Houdré, T. Stoyanov, Expansion and isoperimetric constants for product graphs, http://www.math.gatech.edu/~houdre/research/preprints.html.
[7] C. Houdré, P. Tetali, Isoperimetric invariants for product Markov chains and graph products, http://www.math.gatech.edu/~houdre/research/preprints.html.
[8] M.R. Jerrum, A.J. Sinclair, Conductance and the rapid mixing property for Markov chains: the approximation of the permanent resolved, in: Proc. 20th Annual ACM Symposium on Theory of Computing, 1988, pp. 235–243.
[9] R. Kannan, Rapid mixing in Markov chains, ICM III (2002) 673–683.
[10] J.G. Kemeny, J.L. Snell, Finite Markov Chains, Springer-Verlag, New York, 1983.
[11] L. Lovász, M. Simonovits, Mixing rate of Markov chains, an isoperimetric inequality, and computing the volume, in: Proc. 31st Annual Symp. on Found. of Computer Science, IEEE Computer Soc., 1990, pp. 346–355.
[12] L. Lovász, Combinatorial Problems and Exercises, Akadémiai Kiadó, Budapest, 1979, North-Holland, Amsterdam, 1993.
[13] A.J. Sinclair, M.R. Jerrum, Approximate counting, uniform generation and rapidly mixing Markov chains, Inform. and Comput. 82 (1989) 93–133.
[14] F. Spitzer, Principles of Random Walk, Springer-Verlag, New York, 1964.
[15] C. Thomassen, Resistances and currents in infinite electrical networks, J. Combin. Theory Ser. B 49 (1990) 87–102.